A Review of Traditional and Data-Driven Approaches for Disruption Prediction in Different Tokamaks

M. Priyanka¹, J. Sangeetha¹,*, and C. Jayakumar²

¹Department of CSE, Srinivasa Ramanujan Centre, SASTRA University, Kumbakonam, India
²Shakthi Engineering Works, Neyveli-607803, India

Abstract. Tokamak is a nuclear fusion reactor; inside, the two lighter nuclei known as deuterium and tritium are first ionized together to form plasma, which is heated up to 150 million degrees Celsius, and then they are confined by the torus-shaped magnetic field. During this process, it releases a massive amount of energy, making fusion a feasible option for a long-term and renewable source of energy. On the other hand, plasma leads to disruptions as a consequence of the sudden implosion of the system, which halts the fusion process. Disruptions can irrevocably harm current fusion devices and are predicted to have a more catastrophic impact on future devices such as ITER since they cause a rapid loss of confinement. To control, and prevent disruptions, or at least lessen their negative impact by mitigating them, various traditional and data-driven models obtained with machine learning and deep learning techniques have been used, an overview of some of which is presented in this article. These models are commonly used to forecast their occurrence and give sufficient time to take some counteractive measures.

1 Introduction

With the use of many advanced technology and equipment, there are several ways and methods used to produce power and electricity throughout the world. Nuclear reactors are one of the techniques, and their ability to generate electricity is also thought to be more reliable than other sustainable energy sources like solar, wind, etc. As a result, nuclear power plants are extremely efficient and only need small amounts of fuel to create large amounts of power.

The tokamak is one of the nuclear reactors that were first constructed by Russian scientists, and the name “tokamak” is a Russian acronym known as “Toroidal Magnetic Confinement”. But later onwards, it can be referred to as a “Toroidal Chamber with Magnetic Coils”. The fusion reaction [1] is being employed within the reactor in place of the fission approach [2] because of the liberation of highly radioactive nuclear waste [3]. There are several types of tokamak reactors available across the world, and some of them will be briefly explained in the following section of this article.

*Corresponding author: sangita.sudhakar@gmail.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).
Fusion research addresses a wide range of complex technological and physical issues. These issues include plasma stability, energy and particle exhaustion, reactor safety, environmental compatibility, and so on. Among these issues, disruption is the most pressing concern for tokamak reactors. Disruption is defined as the fatal loss of plasma control, which results in the termination of plasma confinement. As a consequence, a massive volume of particles is expelled from the confinement, potentially damaging the machine's structural components and destroying the energy that has been stored. It happens when the tokamak plasma becomes so sparse that the stored thermal and magnetic energy is rapidly lost. Therefore, disruption is considered to be a primary challenge to producing efficient energy during the tokamak operation.

Research on forecasting disruption in tokamak reactors has evolved into a variety of methods that can be generally classified as traditional techniques and data-driven techniques. Between these two methods, data-driven techniques are currently more prevalent than classical techniques due to their ability to manage and preprocess data from multiple diagnostic signals rather than handling the raw data from unique signals and also achieving an accurate prediction result. Previously, disruption prediction was accomplished primarily through the use of traditional methods such as signal processing, discrete wavelet transform, and so on. The main drawback of employing these methods is that prediction is performed by taking only a few signals rather than the variety of signals from throughout the tokamak. The selection process is critical in this case because it has a direct impact on the model's performance.

This article is structured as follows: Section 2 will provide an overview of tokamak architecture and operating principles; Section 3 will review the traditional and data-driven approaches to disruption prediction; Section 4 will reach a key conclusion by discussing several promising research directions for these data-driven techniques.

2 Overview of Tokamak

The tokamak is an experimental nuclear reactor that was mainly designed to harness fusion energy. It uses the energy that comes from the fusion reaction, a process of merging two or more lighter nuclei into a single heavier nucleus. Although some radioactive elements are released during this process, the life span of these elements is exceptionally short, unlike the nuclear fission reaction.

![Diagram of tokamak reactor core](image)
Fig. 1 depicts the outline structure of a nuclear reactor, which explains how tritium and deuterium can be fused to produce a plasma, which is confined by a magnetic field to generate heat energy. This heat energy is converted into steam, which drives the turbine and generates electricity. As a result, the tokamak reactor is chiefly used to generate electricity.

Around the world, different tokamak nuclear reactors are accessible, and only a few of them are presented in this article. They are as follows:

1) Joint European Torus (JET) [6]: It has been the biggest and most sophisticated tokamak in the world for 20 years. It is operated by the Culham Centre for Fusion Energy and is located in Europe. It is constructed in the early 1980s and used by over 40 European Laboratories.

2) EAST (Experimental Advanced Superconducting Tokamak) [7]: It began operations in 2006 and is located in China under the auspices of the Hefei Institutes of Physical Science, Chinese Academy of Sciences. It is more flexible and almost identical in equilibrium and shape to the International Thermonuclear Experimental Reactor (ITER) tokamak.

3) DIII-D [8]: is one of the largest tokamak facilities in the United States, operated by General Atomics in the Department of Fusion Energy. It is commissioned in 1986. The name "DIII-D" was derived from its plasma shape, which resembles the letter "D."

4) GOLEM [9]: A circular-section experimental tokamak controlled by the Faculty of Nuclear Sciences and Physical Engineering. It is run by the Czech Technical University (CTU) and is primarily used for educational purposes for both foreign and home-grown students. The ability to control this device remotely via the internet is one of its features.

5) ADITYA [10] is a medium-sized tokamak that was put into service in 1989. This tokamak is located at the Institute for Plasma Research (IPR) in India and has been used therefore more than 25 years in a circular poloidal ring limiter configuration.

Fig. 2

Other experimental reactors are also listed in Fig. 2, but their parameters are completely distinct from one another. From that ITER is one of the largest fusion reactors in the world, which is intended to serve as a crucial experimental step between current fusion research machines and fusion power plants of the future. For every 50 MW of input heating power, it is intended to produce 500 MW of fusion power. Since it is the first fusion device to
produce net energy, it will go down in history. The ITER tokamak is being built and run in cooperation by 35 nations. However, the process is still ongoing, and it is hoped that full fusion will be accomplished in 2035. The parameter comparison between the tokamak reactors is expressed in Table 1.

### Table 1: Design Parameters for tokamak reactors

<table>
<thead>
<tr>
<th>PARAMETERS</th>
<th>JET</th>
<th>EAST</th>
<th>DIII-D</th>
<th>GOLEM</th>
<th>ADITYA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major Radius (R)</td>
<td>2.96 m</td>
<td>1.75 m</td>
<td>1.74 m</td>
<td>0.4 m</td>
<td>0.75 m</td>
</tr>
<tr>
<td>Minor Radius (a)</td>
<td>2.1 m</td>
<td>0.4 m</td>
<td>0.56 m</td>
<td>0.085 m</td>
<td>0.25 m</td>
</tr>
<tr>
<td>Toroidal Magnetic Field Strength (B_T)</td>
<td>3.45 T</td>
<td>2 T</td>
<td>1.9 T</td>
<td>0.8 T</td>
<td>0.9 T</td>
</tr>
<tr>
<td>Plasma Current (I_P)</td>
<td>3.2 MA</td>
<td>0.5 MA</td>
<td>1.1 MA</td>
<td>0.025 MA</td>
<td>0.08 MA</td>
</tr>
</tbody>
</table>

#### 2.1 Architecture and its working principle

As seen in Fig. 3 (Ref link: https://www.iter.org/proj/inafewlines), a tokamak is a ring-like doughnut-shaped vessel with magnetic coils encircling it to provide a trap for the fusion fuel. Deuterium and tritium fuse together inside the tokamak, and as this fuel is heated to more than 100 million degrees Celsius, the electrons are separated from the particles and produce “plasma” [11]. Plasma is a gas of ionized atoms and electrons present inside the chamber known as the plasma chamber. Plasma particles begin to heat up as they become energized and collide. Auxiliary heating methods aid in raising the plasma temperature to fusion temperatures between 150 and 300 million degrees Celsius. When “energized” particles collide, they can overcome their natural electromagnetic repulsion and fuse, releasing massive amounts of energy.
power a steam engine and generate electricity. Divertor plates permit the removal of waste from the plasma in real-time while the reactor is still running. The feeders transport the electrical power and cryogenic supply to the ITER magnets via the warm-cold barrier.

To confine the plasma, three sets of magnetic coils are used. A powerful "Toroidal" field is produced by one set of magnetic coils and is directed circumferentially around the torus. A second magnetic field is produced by a central Solenoid, an electric current-carrying magnet, along the "Poloidal" direction, which is the shortest path around the torus. The two field elements produce a tangled magnetic field that holds the plasma's particles in place. An outer poloidal field created by a third set of field coils molds and arranges the plasma.

3 Disruption predictions

When tokamak plasma instability becomes severe enough to quickly deplete the stored magnetic and thermal energy, a disruption takes place. In order to promote effective tokamak reactors, disruption prediction, avoidance, and mitigation are crucial. Typically, there are two kinds of approaches to solving this issue: i) Traditional Techniques, and ii) Data-Driven Techniques.

3.1 Traditional Techniques

To understand the plasma activity, a study of current perturbations inside the tokamak is required. Only magnetic field perturbations can be used to control high-beta poloidal Magnetohydrodynamic (MHD) activity and major disruptions. Mirnov signals are recorded by the Mirnov coil for current perturbation studies where the helical nature of the perturbed current has no effect on the mirnov signals. Green's function is used to analyze the relationship between current perturbations and mirnov signals on a rational surface in tokamaks. The shift of the plasma column in the vacuum chamber, satellite formation, and phase modulation effects can all be studied using mirnov coil signals [12].

For MHD-based disruption prediction, a traditional method based on the signal processing technique [13] has been offered. To recognize non-stationary signals for time-frequency analysis, a Choi-Williams distribution is used. For this investigation, the researchers utilized data from the Electron Cyclotron Emission (ECE) experiment and magnetic pick-up coil signals. This method allows for the identification of MHD instabilities such as Neoclassical Tearing Modes (NTM), Edge Localized Modes (ELM), and saw tooth crashes between ELM and Wash Board (WB) procedures. To demonstrate its superiority over others, the proposed technique is compared to spectrograms and the Wigner distribution. A Choi-Williams distribution, as opposed to a spectrogram, provides a more accurate picture of the interaction between ELM and WB. By selecting the value judiciously, this technique also provided a high time-frequency resolution and artifacts. A Choi-Williams distribution is compared to wavelet approaches for analyzing non-stationary fusion plasma signals. To improve the short-time Fourier transform spectrogram for analyzing nonstationary signals in plasma, the Choi-Williams distribution and continuous wavelet transform scalogram were used. A comparison study was conducted to demonstrate the advantages of the Choi-Williams distribution over wavelets for improving spectrogram presentation. The real-world signals of magnetic pick-up coil signals are analyzed using spectrograms, scalograms, and Choi-Williams. When removing artifacts, the Choi-Williams distribution outperforms the spectrogram and scalogram techniques [14].

Understanding the dynamics of plasma turbulence and the driving mechanisms in a tokamak is a major challenge. The plasma turbulence, as well as the MHD activity in the ISTTOK tokamak, are both examined [15] using an Empirical Mode Decomposition (EMD) method. Langmuir probes were used to analyze edge plasma density, and mirnov coil signals were used to analyze perturbed magnetic signals caused by current.
MHD activity occurs in a rapid burst in the ISTTOK tokamak due to the fast time scale, where the plasma pulse length is 30 ms and particle confinement is on the order of 0.3 ms. As a result, EMD is used to identify the time-frequency components of spatial modes. The EMD method revealed that the increase in turbulence spectra frequency is related to biasing and concurrence with MHD activity. EMD is also used to investigate floating oscillations, revealing intrinsic mode functions that provide a clear understanding of MHD oscillations.

To analyze the mirnov coil and edge fluctuation data from ADITYA tokamak, the authors suggested data analysis techniques [16] based on EMD and Hilbert transform. The scrape-off layer plasma is measured using Langmuir probes and magnetic oscillations through mirnov coils during an EMD procedure. When applied to signals, the EMD technique produces edge fluctuations with a finite number of 10 modes. The strongly correlated modes have a frequency range of 10–60 kHz from raw data, and signals with a frequency of 20 kHz are almost stationary with low-frequency components. The intermittent modes in Aditya tokamak data have a bandwidth of 30 kHz and are composed of non-stationary high-frequency components. The presence of high-frequency components is explained by fluctuations in plasma pressure and diamagnetic frequency. This technique is further used to study the three-mode interactions and triplet interactions among high-frequency modes on fluctuation data.

To identify the magnetic perturbations in a toroidal plasma, a Fourier decomposition with a singular value decomposition technique has been proposed [17]. In this article, Fourier time and spatial basis functions were used to find the least residue, and Singular Value Decomposition (SVD) to find the best fit on a subset of signals. On the mirnov coil data, the proposed technique yields the best fit of mode amplitudes, mode numbers, and phase modulations. The Eigenmodes in toroidally confined plasma are identified using this method, where weak magnetic signals with longer coherence times are resolved. The proposed method was applied to the outboard mirnov coil signals of MAST tokamaks to analyze the high-frequency components of Alfvénic activity and low-frequency activities.

A method for estimating poloidal wavenumbers for magnetic fluctuations in the ISTTOK tokamak was put forth [18]. Determining the wavenumber and amplitude of signals for tracing plasma instabilities becomes a priority for predicting disruptions. In this article, unstable modes are observed using wavenumber spectra and a pair of poloidal mirnov coils. A real-time investigation of the poloidal mode number is estimated using Kalman filtering and cross-spectrum analysis. Kalman filtering was used rather than the traditional Fourier transform. The proposed technique was tested on shot no. 17081 of the ISTTOK tokamak, which clearly shows that mode m=2 rotates at 160 kHz in the direction of electron diamagnetic.

For MHD structure analysis in FTU tokamak or disruption avoidance through ECRH, a singular value decomposition technique was used [19]. An offline-based mode analysis on mirnov coil signals and tomography on (Soft X-Ray) SXR signals is performed. The researchers used SVD on the same shot as FFT to determine the best fit for mode identification. The shorter execution time and smaller interval presented preliminary results indicating that SVD is better at identifying mode numbers for MHD-based disruption prediction.

The SVD technique was suggested as a mode number identification [20] for FTU tokamak mirnov coil signals. To extract the poloidal m and toroidal n mode numbers from mirnov coil data, the SVD technique identifies the principal components and principal axes. The mode numbers aid in identifying tokamak instabilities. To identify the exact phase changes in plasma, the poloidal and toroidal mode numbers must be compared with mode location by reconstructing the q-profile.

Identification in the DIII-D tokamak MHD mode is done with mirnov coil signals.
Understanding the MHD behavior of plasma requires the identification of coherent waves from unstable plasma. Mode identification from conventional poloidal and toroidal mode numbers has become complex and difficult due to plasma shaping, uneven distances between detectors and resonant surfaces, and toroidicity. In this proposed technique, the directions of maximum coherence in fluctuation data are identified with their own basis vectors using SVD [21]. The spatial structure of modes was provided by the spatially distributed Eigenvector on time and the covariance matrix with energy content. The SVD approach reduces noise and significant basis vectors while minimizing storage space for fluctuated data restoration. The proposed method identifies mode numbers for MHD-based disruption prediction, as well as the evolution of coherent structures and energy contents of modes in plasma.

For a JET tokamak, the researchers suggested a time-frequency and mode number estimation [22] for MHD analysis using a Kalman filtering approach. When the new-classical tearing mode transitions to the mode-locking mode, there is a risk of disruption as well as mechanical damage to tokamak components. The authors used the Kalman Filter (KF) in this study to estimate the amplitude and phase evolution of tokamak instabilities with predefined mode number declarations. Two KFs have been used: a non-linear Kalman filter serves to isolate the dominant non-stationary signals, and a linear Kalman filter identifies the mode number based on predefined mode number declarations. The authors used real-time JET tokamak data to present their study using Kalman filtering.

In the IRT-1 tokamak, the authors conducted a comparison study of Fourier analysis and SVD for mode number identification [23]. According to the experimental results, the authors claim that when the energy between modes is balanced, the Fourier transform is unable to recognize the modes. According to the experimental results, SVD is more precise than the Fourier series. When the energy distribution is proportional and the mode numbers are high, the Fourier analysis determines incorrect mode numbers. According to the results of this survey, the Fourier method fails for high-frequency mode numbers and superposition of mode numbers.

Using Choi-Williams Distribution (CWD), the authors proposed a time-frequency analysis [24] for non-stationary MHD signals. In order to pinpoint the rapid MHD modes and events occurring in plasma, the data from the HT-7 and EAST tokamaks, mirnov coils, and SXR have been used. A comparison of the Choi-William over continuous wavelet scalogram and STFT was performed. According to the experimental results, CWD will behave like Wigner distribution with optimal frequency resolution if and smaller artifacts reduce the resolution of the extracted signal. The proposed method detects the emergence of high-frequency saw tooth, tearing, and low-frequency MHD instabilities in ELM.

Using mirnov coil signals, the researchers carried out an FFT-based MHD analysis [25, 26]. To identify modes in a tokamak, poloidal Mirnov coils are used. When FFT is applied to mirnov coil signals, the power spectral density is plotted to estimate the signal's power with respect to frequency. The edge safety factor was estimated using a power spectrum up to a frequency limit of 33 kHz, and the mirnov coil data was plotted using power spectrum density in the presence of an external field and in the absence of a field. A technique known as SVD-based mode number identification from mirnov coil signals has also been performed. When compared to the FFT technique, the SVD technique was found to extract mode numbers very efficiently, even in mode coupling and mode separation, which is best for the IRT-1 tokamak [27].

Based on time-frequency analysis of mirnov coil signals in the ADITYA tokamak, the authors proposed a novel Hilbert-Huang Transform (HHT) [28, 29]. The HHT consists of two major processes: a shifting process that extracts a finite number of mono-frequency components as modes or intrinsic mode functions, and applying the Hilbert Transform on extracted IMFs to extract instantaneous frequencies and amplitudes. In comparison to the S-ES Web of Conferences 477, 00039 (2024) abstract

STAR 2023
Fourier transform, the HHT can be used directly on non-stationary signals, and it outperforms wavelet techniques as well.

HHT and MHD mode identification were combined in an SVD technique that was used to identify the Principal Axes (PA) and Principal Components (PC). The PAs are the dominant harmonics in signals that are processed further using the EMD approach. IMFs are identified as time-frequency behavior and plotted to extract frequency components at the corresponding time period. The spatial and temporal analysis of mirnov coil data is performed using this method to identify mode numbers and dominant MHD frequencies. In this work, the authors have been using the mirnov coil data from the IRT-1 and Golem tokamaks.

3.1.1 Discussion of Traditional Methods

Previously, anticipating disruption on the tokamak was frequently done using traditional approaches. Even though these techniques are simple to create and apply, it is challenging to foresee plasma disruption in real-time when samples are taken from several signals. The model is trained using only the raw data from the distinct signal in these techniques. Assume that if the data pattern from the different signals is comparable, then the prediction won't be impacted. However, this is not possible during the real-time testing phase because the time series data for different diagnostic signals varies slightly. It should be noted that using raw data without any preprocessing or normalization is the main issue that could compromise the accuracy of predictions. As a result, some alterations to the traditional methods are needed to enhance the performance of the model, such as their integration with other methods.

3.2 Data–Driven Techniques

On the basis of a number of datasets, data-driven models are built. The performance of the model will be justified using the available or past information as standard data, and it will also help to construct the model's algorithm. To be more precise, the data-driven model parameters will be meticulously chosen and altered through methodical comparison between the available information and the output of the model. Suppose, if the disruption was foreseen by the model, then a real-time predictor tool based on these methodologies would give adequate time to take preventative or mitigating action.

The following subsection will survey the different algorithms based on Machine Learning (ML) and Deep Learning (DL) approaches that have been used on various experimental tokamak reactors.

3.2.1 JET (Joint European Torus)

JET is the fusion reactor that uses a magnetic field to confine the hot plasma at a temperature of 150 million Kelvin, which is normally 10 times hotter than the sun. The main parameters of this reactor are shown in Table 1. The large amounts of energy particles are released due to the sudden loss of plasma confinement, which produces heavy heat loads and electromagnetic forces, causing damage to the machine itself. Even though the root cause of the disruption is not completely understood, various types of classifier techniques or predictor tools based on ML/DL approaches help to predict the disruption and classify them. The following segment discusses a few of the data-driven approaches based on these strategies:

3.2.1.1 Support Vector Machine (SVM)

Over the past year, disruption has been predicted using the black box method. The drawback of this strategy is that if the system is not updated, it may behave degenerately, which means that the new plasma's features will be different from those of the observed plasma that was used for training. It may result in a misclassification problem; to address this, novelty detection and prediction were performed using Support Vector Machines [31]. The key disadvantage of novelty detection was the difficulty in labeling unfamiliar data points because it was unaware of them throughout the training phase. A composite
impending disruption indicator, also known as a warning indicator, has been provided by combining the multiple plasma diagnostic signals. During the online application testing phase, Decision Function Value (DFV) obtained valuable information about the predictor output’s trustworthiness and the novelty of input. The predictor in this approach could predict novel samples based on DFV. During the testing phase, if the DFV value lies within these bands $[1-\Delta a, 1+\Delta b]$ and $[-1+\Delta a, -1-\Delta b]$, then the given sample is not novel; otherwise, it is a novel sample. The following Fig. 4 depicts the SVM working procedure based on the DFV value.

Fig. 4. Working flow of SVM based on DFV value

The optimization approach has been used to choose the percentage of data samples for training purposes. It helps to avoid overfitting problems in the training phase as well as reduce the errors in the cross-validation set, which results in an optimal value of 90%. Thus, the robustness and performance of this system were increased because it was able to identify the new incoming pulses based on its experience in the training phase. Although it can solve non-linear classification problems, its discrimination capability was reduced when the performance of the system increases.

3.2.1.2 Different Machine Learning Tools

To mitigate or avoid disruptions, an alarm system has been designed by using various machine learning techniques [32]. There were two main reasons for this approach. First, the phenomenon of identifying the root cause of the disruptions was not possible. Second, there is an enormous amount of disruption data in experimentation. In recent days, predictors have been implemented for the complex system to find out the relationship between the boundaries of operational space and the plasma state, which would be free from disruption by using avoidance actions instead of mitigation. So, these kinds of predictors took suitable avoidance actions by detecting what kind of disruptors were going to occur. To achieve these, two models have been implemented: SOM (Self-Organizing Map) [33] and GTM (Generative Topographic Mapping) [34, 35]. For the training set, before taking the diagnostic signals as an input, it was necessary to normalize all the data because the signal range might vary in magnitude. The evaluation of the prediction performance has been visualized in the following Fig. 5.
3.2.1.3 Neural Network

Deep learning was mainly used for processing the huge amount of data that was collected from the fusion diagnostic system. The different types of plasma disruption have been classified by using the neural network [37]. It takes input in the form of various diagnostic signals to classify disrupted pulses. In an earlier stage, the classification process was done manually, but it was a very tedious and time-consuming process. To overcome this problem, multiple classifier techniques have been used based on MLP. It consists of 10 diagnostic signals, 7 disruption classes, and 149 disruption pulses. For every pulse, it takes three consecutive samples as an input sequence and the sampling time was 40ms. Among the seven disruption classes, the Vertical Displacement (VD) class was excluded because its time frame excludes its classification at a disruption time of 40 milliseconds. While on the other hand, the Horizontal Displacement (HD) class was included in the training set for further training. However, as demonstrated in MLP, mitigation only predicts the occurrence of the disruptive event without knowing its root cause. To resolve this, the GTM method has been used. It mainly focuses on identifying the plasma regions in 2D mapping. In this map, each node was considered as a sample that comes from various kinds of non-disruptive and disruptive samples. This map could be suggested to act as a classifier as well as a predictor. Throughout this map, the following two actions would take place: (i) The first action was that each sample could be related to its suitable class (disruptive or non-disruptive). (ii) The second action was it triggers the alarm when it detects disruptions. As a result, the overall prediction's success rate has been achieved at 89%.
other hand, the conditions for Density Limit (DL) and Radiated Power (RP) classes are similar, so those classes have been merged into one. As a result, the 7 disruption classes have been reduced to 4 classes. The training dataset has 109 pulses, out of which 20% of pulses were assigned for cross-validation, and the testing datasets have 20 pulses. Almost the majority of the network produces a similar result, however, at the time of choosing the best classifier, it led to a loss of potential or valuable information. To resolve this, a combined approach could be used within the classifier technique. One was majority voting, and the other one was averaging the output space [38]. So, the experiment was conducted based on these approaches, and the result shows, that the performance of the neural network is considerably high and also that it is a very reliable, fast, and feasible technique for classifying the disruptions automatically.

3.2.1.4 Artificial Neural Network (ANN)

For further improvement, the disruption prediction process has been mainly focused on flat-top disruptions [39]. During the flat-top scenarios, the shape, steady equilibrium position, and quasi-stationary plasma current have been obtained, and this process was successfully done only by controlling the plasma. The discharge selection process has been completed by satisfying the three criteria: (i) $I_{\text{pl}}$ (Plasma Current) exceeds 1.5 MA, (ii) $X$-point configuration, and (iii) flat-top plasma current profile [40]. For training and testing purposes, it took a database consisting of totally 102 safe pulses and 92 disrupted pulses. For every pulse, 40 samples were taken as an input sequence, and the sampling time was 20 ms. The value of $t_\alpha$ was equal to 320 ms for DL (Density Lock) disruption, and $t_\alpha$ was equal to 28 ms for other disruption classes. Here, 71 successful pulses and 66 disruption pulses have been assigned to the training dataset, 15 successful pulses and 11 disruption pulses were allocated to the cross-validation set, and the test dataset contains 16 successful pulses and 15 disruption pulses. The Artificial Neural Network has been used for training purposes, and the disruption predictions have been done in 100 ms before $t_D$ (Disruption time). Even though it is a reliable method, the percentage of false alarms and missing alarm rates were high.

3.2.1.5 Deep Neural Network

The neural network takes the input in various forms, such as temporal data, and images. Two models [41] have been popularly used to handle this kind of data. The first one was the Convolutional Neural Network (CNN), which processed the images for plasma tomography, and the second one was the Recurrent Neural Network (RNN), which predicted the disruptions by analyzing the time series. In the CNN model, the plasma state would be monitored by using the specified signals that were derived from the diagnostic bolometer system [42], and it would detect the accumulation that occurred at the plasma core and also the impurity transportation. It is the basis for reconstructing the plasma tomography that gives the plasma radiation profile in 2D form. However, it required more time to reconstruct itself. To overcome this problem, CNN has been trained by using a wide collection of tomography samples. There were two main principles behind CNN for image classification [43]. They are: (i) In the first stage of the layer, the useful pieces of information have been extracted from the given input image. (ii) In the second stage, the classification process has been performed based on these extracted features. In the RNN model, the data from these specified trained signals provide useful information to predict the disruptions in terms of both the remaining time for an impending disruption and the probability of disruption. During this prediction, the accuracy was increased by using various parameters, which have to be considered as an additional input to the RNN neural network. The main advantage of using this approach is that the relevant features have been directly extracted from the raw signals. As a future scope of this model, the researchers considered that data should be extracted from multiple diagnostic systems rather than just one to implement a large deep learning model.
3.2.1.6 Increased Time Resolution and Robustness of JET APODIS

APODIS (Advance Predictor Of DISruption) is a real-time disruption predictor that has been developed for the JET tokamak. It has been used for seven plasma quantities, such as plasma internal inductance, diamagnetic energy time derivative, mode lock, high radiated power, plasma current, input power, and plasma density. However, for any discharge, there will be a signal failure, which leads to producing some incorrect signals. Those signals could be utilized by APODIS as an input, which would cause some major issues in predictions. To overcome this problem, this proposed article [44] presented two methods. They are as follows: (i) simulating anomalous signals; (ii) sliding window mechanisms.

First, the anomalous signals were simulated to determine the reliability of the predictor system, which helps to estimate the quality of APODIS. The mode lock and plasma inductance signals are more essential for APODIS than the other six plasma signals (plasma density, plasma current, plasma inductance, mode lock, radiated power, input power, and stored diamagnetic energy). It indicates that the success rate was slow when any failure occurred in these two signals. However, the failure that occurred in other signals has less impact on the success rate.

Next, the sliding window mechanism has been used to improve the temporal resolution of the disruption predictor. Due to the lack of this resolution [45], some disruption might be missed. The important note is that the predictor was always enabled only when the plasma current was above the threshold value [750kA]. Once it crossed the threshold values, there was no prediction. It occurred because the predictor was disabled, which led to the missed disruption. This mechanism aimed to change the resolution but not the sampling rate. By using this window mechanism, the implementation of APODIS has been carried out in different time resolutions (1, 2, 4, 8, and 16 ms) to trigger alarms rather than the current resolution of 32 ms. By increasing the temporal resolution, a better success rate (83%) and warning times have been achieved and the false alarm rate was also increased.

3.2.1.7 Genetic Algorithm for Real-Time Disruption Prediction

For advanced development, the performance of the predictor system has been improved by choosing a specific set of parameters [46] for the input signals and it could be done by extracting the features based on the Genetic Algorithm (GA). Comparing the proposed result with their existing result would show the improvement in success rate and also that the time interval has been extended before the disruption occurred. The significance of a signal can be firmly impacted by its two kinds of representation. The first one is properties (Standard Deviation, Mean, ad-hoc transformation), while the second one is the frequency or temporal domain. SVM (Support Vector Machine) is the basic working principle of APODIS, and its architecture has been visualized in the following Fig. 6.

Fig. 6 Architectural view of APODIS
It consists of two layers: (i) The first layer contains the chain of different SVM predictor modules, and the training process was done in parallel. (ii) The next layer contains the Decision Function, which determines whether the alarm is to be triggered or not based on Decision Function Value (DFV). The overall working procedure of this architecture has been carried out by giving three consecutive time windows for an analyzed shot as an input to the first layer. It computes the three values as an output of the first layer. Next, the second layer takes these three values as input to the decision function and processes them. The most important thing is that the first layer process has been done by using disruptive and non-disruptive plasma discharge features, and it could take a calculation time interval of 30 ms. Hence, the alarm had been triggered around 200 ms before the occurrence of the disruption.

Table 2 provides the summary of data-driven approach for JET tokamak.

<table>
<thead>
<tr>
<th>Tokamak Reactor</th>
<th>Data-Driven approaches</th>
<th>Remarks</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>JET</td>
<td>Support Vector Machine</td>
<td>Non-linear classification problems are solved, but the system's capacity for discrimination is diminished</td>
<td>[31]</td>
</tr>
<tr>
<td></td>
<td>Self-Organizing Map, Generative Topographic Mapping</td>
<td>In the current analysis, premature detections are not taken into account</td>
<td>[32]</td>
</tr>
<tr>
<td></td>
<td>Neural Network (MLP)</td>
<td>There were no processed data used</td>
<td>[37]</td>
</tr>
<tr>
<td></td>
<td>Artificial Neural Network</td>
<td>It is reliable, but the false alarms and missed alarms are high</td>
<td>[39]</td>
</tr>
<tr>
<td></td>
<td>Convolutional Neural Network, Recurrent Neural Network</td>
<td>Only a bolometer signal was used. Multiple diagnostic systems were employed to enhance performance</td>
<td>[41]</td>
</tr>
<tr>
<td></td>
<td>APODIS – Simulating anomalous signals, Sliding window mechanism</td>
<td>The false alarm rate was increased along with the success rate.</td>
<td>[44]</td>
</tr>
<tr>
<td></td>
<td>Feature selection based on genetic algorithm</td>
<td>Future research should conduct a thorough physical interpretation to ascertain the causes of the higher performances using this collection of plasma properties in order to provide a reliable solution for subsequent devices.</td>
<td>[46]</td>
</tr>
</tbody>
</table>

3.2.2 EAST (Experimental Advanced Superconducting Tokamak)

3.2.2.1 Long Short-Term Memory for predicting the density limit disruptions
Disruption predictions have been done using the Long Short-Term Memory (LSTM) algorithm because it works very well in the case of time-series data. However, it took a very long time to train the model for the entire flat-top phase. To resolve this problem, the LSTM approach [47] based on the short time sequence has been implemented for an EAST tokamak. This approach helps to minimize the computational time as well as to fade the gradient problem that occurred in the model when it was trained by the RNN.[4] The optimal predictions have been made by comparing the performance of the following models:

(i) Model 1: Training and testing have been carried out directly on the flat-top phase. It took 11 diagnostic signals, 2921 non-disruptive shots, and 300 disruptive shots as input. Among them, 150 non-disruptive shots and 150 disruptive shots have been issued for training purposes, and the remaining shots are for testing purposes. Here, the disruption phase was set to 400 ms to predict disruption in advance. The AUC (Area Under the Receiver Operating Characteristic (ROC) Curve) has been used to estimate the quality of models based on two conditions: (i) AUC = 0.5> classification ability was not good, and (ii) AUC = 1> classification performance was better. In this model, the higher AUC has been achieved at 0.8646, which is nearly equal to 1.

(ii) Model 2: Predicting disruptions based on short-time series: Here, 3221 plasma discharges were used for testing (2921 shots) and training (300 shots) purposes. In the previous model, it took 6900 s per epoch for training. However, in this model, only 36 s per epoch have been taken. The higher AUC would be increased to 0.9379. This experiment shows that better performance was achieved with higher AUC only by training the data based on a short time sequence, and testing has been done on the whole flat-top phase. One of the advantages is that it required less time for training and its performance was improved, but early warnings would result in a significant loss of experimental time in these discharges.

3.2.2.2 Fully Convolutional Neural Network (CNN)

To improve the performance of predictions, a full Convolutional Neural Network [48] has been implemented on the EAST tokamak to classify the disruptive discharges from the non-disruptive discharges. It takes 14 diagnostic signals as an input and can be divided into time slice sequences within a period of 30 ms. In the case of long-pulse operations, the amount of disruptive data was less than the non-disruptive data. So, in the training process, a random selection process was used to select non-disruptive data, and its FA rate was analyzed. To get the optimized model, some valuable data has been artificially added to the dataset. There are various kinds of disruptions in the EAST, and some of them are given below [48]:

1) Disruption by Impurity Radiation: It causes an inconsistent state in plasma temperature.
2) Disruption by Vertical Displacement Event (VDE): The plasma exceeds the PCS (Plasma Control System) limit; it moves very quickly and collides with the wall.
3) Density Limit disruptions: The density level is increased to reach up to the limit of Greenwald density.
4) Magnetohydrodynamic: When instability occurs in the MHD mode, it leads to disruptions.

In this method, a multi-signal threshold value could be used to make predictions rather than a single signal, because it lowers the accuracy level [49]. It was done by using the t-SNE (t-Distributed Stochastic Neighbor Embedding) algorithm and its pieces of information have been extracted from the segment by using one-dimensional CNN. The performance results are shown in Table 3. As a future scope, further improvements have been made to the model by adding cross-device performance and more disruption data on related signals.
Table 3: Resultant Comparison Table

<table>
<thead>
<tr>
<th></th>
<th>Before adding artificial data</th>
<th>After adding artificial data</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPR</td>
<td>0.82</td>
<td>Increased to 0.875</td>
</tr>
<tr>
<td>FPR</td>
<td>6.7</td>
<td>Decreased to 0.061</td>
</tr>
</tbody>
</table>

TPR = True Positive Rate, FPR = False Positive Rate

3.2.2.3 LSTM disruption predictor

The disruption on EAST tokamak has been predicted by implementing the LSTM model [50] and the results were compared with the previously proposed CNN. For this purpose, the same dataset and diagnostic signals (14 signals) that were already used in CNN have been considered as inputs for this model in the training and testing phase. The detailed descriptions of the signals have been shown [48]. Although the prediction performance of LSTM was better than CNN, the prediction time of this model was 14 ms earlier than CNN, which gave the result an increasing rate of FPR. The FPR rate has been reduced by using additional data, such as extracting features from specified signals like the bolometer, which were considered as an input to the model. As a result, the AUC value of the improved LSTM model has been achieved at 0.89 and the false alarm rate was also decreased to 0.094. And also, the prediction time was 8.7 ms earlier than the previous LSTM model. The future work of this model mainly focuses on reducing the inconsistent state of the plasma by adding more specified signals and also analyzing hybrid algorithms like the combination of CNN and LSTM.

Table 4: Summary of data-driven techniques for EAST tokamak.

<table>
<thead>
<tr>
<th>Tokamak Reactor</th>
<th>Data-Driven approaches</th>
<th>Remarks</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAST</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.2.3 Disruption Prediction using Random Forest (DPRF)

One of the classification algorithms known as the Random Forest [52] has been used to make explainable predictions for more than 900 plasma discharges and it was the first time the interpretation of predictions has been done through feature contribution analysis, which before adding artificial data After adding artificial data

AUC | 0.9 | 0.92 |
TPR | 0.82 | Increased to 0.875 |
FPR | 6.7 | Decreased to 0.061 |
helps to build a bridge between the underlying physics and the data-driven model. For various input variables in DPRF, the following observations have been made: (i) If $t > 300$ ms $\rightarrow$ far from disruption, (ii) If $t \leq 300$ ms $\rightarrow$ close to disruption (iii) Otherwise $\rightarrow$ Non-disruptive. In this method, the accuracy rate for predicting the non-disruptive shots was very high. As a future scope, the rectified version of DPRF was required to identify the disruptive shots by using an avoidance scheme that was executed by the controller.

3.2.3.2 Decision Tree and Ensemble Methods

For better predictions, a real-time algorithm has been conducted on the DIII-D tokamak by using various machine learning approaches [53] such as the Decision Tree, and the ensemble methods (Bagging, Random Forest, Extremely Randomized Trees, and AdaBoost). In the Decision Tree approach, the output values were predicted by learning decision rules that were received from the training set. There are two types of decision trees: (i) classification—the output must be discrete; (ii) regression—the output must be real or continuous. In this algorithm, a regression tree [53] has been used to predict whether the output value was either occurring in tearing mode or disruptive. In the ensemble approach, four different methods have been used: (i) Bagging—It makes various copies of an enormous number of data points from the original dataset. Then the Boostrap [54] procedure was applied to each copy to randomly choose data for building a new dataset. The most important point was that the size of the elements in the new set was the same as in the original set. (ii) Random Forest—It is almost similar to the bagging model. The best splitting process has been done by randomly choosing a subset of features rather than all after computing the new data set. (iii) Extremely Randomized Tree—Unlike the random forest, the splitting process has been done in a randomized manner [55]. It also provides a different set of splits randomly, and then the splitting process of a node is done by picking the best split among those random ones. (iv) AdaBoost—The accuracy [56] has been improved for the base algorithm by using this method. At each iteration, the set of weights has been modified. Unless it was increased, the data would be predicted incorrectly; otherwise, the value would be decreased.

These ensemble method stake input as a wide range of signals and calculate its variance and mean within a specific period. In addition to this, the algorithm has been developed in two scenarios: (i) Ramp down Control: In this scenario, the regression tree algorithm has been used to predict the disruptions. It receives signals from rtEFIT (real-time diagnostics). As long as the level of disruption was increased to a level that was higher than the threshold value, it indicates some abnormal discharges we are going to occur. Using Off-Normal Fault Response [ONFR] [57], a fast ramp down was triggered by this event, and the plasma termination had been done within 200 ms. (ii) Tearing Mode Avoidance: It was similar to the disruption predictor, and it was mainly used to keep the plasma boundary stable. The tearing mode database has been used for training and testing purposes, and the results are mentioned in Table 5.

From the above methods, the primary benefit of using a decision tree was that it provided the critical relationship [58] for each input variable. However, while using bagging models, some of the elements in the original dataset have not appeared in the new dataset and also it contains repeated data.

<table>
<thead>
<tr>
<th>Threshold Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t = 0.16$</td>
<td>Tearing mode was 90% and false-positive was 10%</td>
</tr>
<tr>
<td>$t$ value was very high</td>
<td>The true positive rate was decreased along with the median. Alarm rate and false-positive rate.</td>
</tr>
<tr>
<td>$t &gt; 0.35$</td>
<td>False positives were less than 3%, and the TPR was achieved at 63%.</td>
</tr>
</tbody>
</table>
3.2.3.3 Temporal Convolutional Network (TCN)

An advanced architecture based on dilated convolutions known as the Deep Convolutional Neural Network (DCNN) [59] has been developed for predicting plasma disruptions. Unlike the RNN/LSTM network, it can track the long-range dependencies in the sequences. The prediction has been done by taking raw data as input from a single diagnostic signal known as an ECEi (Electron cyclotron Emission imaging) [60]. The main role of ECEi was to measure the quantity of the plasma, such as the temperature of the electron, with high temporal and spatial resolution. This kind of prediction problem should be treated as a binary classification problem. And the input was given to the Temporal Convolutional Neural network with four hidden layers to train the model. If the time slice was less than 300 ms, then it could be labeled as a “disruptive” class; otherwise, it was categorized as a “non-disruptive” class [61]. The overall accuracy was achieved at approximately 94%. But, due to their imbalanced classes, the F1 scores give better results, and they have reached approximately 91%. As a future scope, combining various models by including more signals as an input rather than one, to predict the disruptions earlier.

Table 6: Overview of data-driven methods for DIII-D tokamak.

Table 6:

<table>
<thead>
<tr>
<th>Tokamak Reactor</th>
<th>Data-Driven approaches</th>
<th>Remarks</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIII-D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Random Forest</td>
<td>Non-disruptive shots were predicted far more accurately than disruptive shots.</td>
<td>[52]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ML approach – Decision Tree Ensemble methods – Bagging, Random Forest, Extremely Randomized Trees, AdaBoost</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temporal Convolutional Network</td>
<td>It uses information from just one signal, called ECEi. More signals should be incorporated in the future to enhance model performance.</td>
<td>[59]</td>
<td></td>
</tr>
</tbody>
</table>

3.2.4.1 Classifier Techniques

When employing the bagging algorithm, the updated dataset contains repeated data and some of the data from the original datasets might not be there. [53]
3.2.4.2 Ensemble Learning Methods

The existing machine learning model becomes more vulnerable to identifying disruptions. To rectify this, an ensemble-based machine learning approach embedded with the active learning technique has been proposed for predicting disruption. The ensemble model consists of logistic regression, CAT Boost, and REPTree techniques for identifying the class probabilities of given instances. The active learning approach incorporated into the model identifies the class labels for the class probabilities, and the model is efficient in the case of unlabeled data as well. The proposed model utilizes 14 diagnostic features with 117 normal shots and 70 disruptive. The model was found to have an overall accuracy of 98.75%. Still, the model lags in identifying the exact timestamp of the disruption window for the mitigation process in the GOLEM tokamak.

3.2.4.3 Stacking approach

For the next level, a supervised ensemble-based machine learning approach was proposed for disruption prediction in the GOLEM tokamak. In this methodology, the researchers utilized 9 diagnostic features for disruption prediction, consisting of 117 normal and 70 disruptive shots. The proposed model consists of a level-1 based stacking approach utilizing 3 REPTree classifiers and 1 linear regression classifier. The REPTree classifier reduces the error caused by tree growth by pruning the unwanted or less informative carrier branches. Other tree-based classifiers like CART, C4.5, and Random Forest suffer from tree growth and branches, which increases the computational time and maximizes the error rate. The proposed model consists of three REPTree classifiers, which is an advanced version of C4.5 working on the principle of split rule or gain ratio. The proposed technique performed well with an overall accuracy of 97.3%, a ROC of 0.9962, and a Precision-Recall Convergence of 0.9961. Although it has some disadvantages that the proposed model was designed based on homogenous classifiers, so it increases the bias and variance of the overall approach.

3.2.4.4 Kernelized Support Vector Regressive Machine-based Modified Variational Mode Decomposition (KSVRMMVMD) Method

For better performance, Magneto-Hydrodynamics (MHD) based disruption prediction has been implemented on GOLEM tokamak. The Mirnov coil signals are utilized for capturing the MHD mode in tokamaks for understanding the plasma status. A combination of signal processing-based machine learning approaches is utilized for identifying the underlying time-frequency changes in non-stationary signals. Such as the modified version of the Variational Mode Decomposition algorithm is proposed as KSVRMMVMD. This proposed technique identifies the intermittent bands present in the nonlinear signals known as intrinsic mode functions, which were extracted by using the VMD algorithm. Unlike the other existing signal processing algorithms, this proposed model was found to be with an overall accuracy improvement of 19% and with a lower computational time of 40ms.

Table 7 summarized the data-driven technique for GOLEM tokamak.
3.2.5 ADITYA TOKAMAK

ADITYA is the first Indigenous tokamak which was commissioned in the year 1989 [10]. A huge number of experiments have been conducted by taking more than 30,000 plasma discharges with plasma current whose range lies between 50 KA and 150 KA. The key parameters of this reactor are mentioned in Table 1. Disruptions are a major problem in the tokamak, causing structural damage and component failure on occasion. In order to predict disruptions in the ADITYA tokamak reactor, a variety of neural network employed; the most noteworthy works are covered in this section:

3.2.5.1 Artificial Neural Network

An Artificial Neural Network [71] based disruption predictor has been used to predict the density limit disruption [72] in advance. In this approach, the disruption has been avoided by controlling the plasma density within the limit. Two aspects of aims were primarily considered for prediction, which is shown in Table 8:

<table>
<thead>
<tr>
<th>Point Of View</th>
<th>Has been done</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boundary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prediction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prediction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subset</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decision</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controllers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mitigation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disruptions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avoidance</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7: Summary of data-driven approaches for GOLEM

<table>
<thead>
<tr>
<th>Tokamak Reactor</th>
<th>Data-Driven approaches</th>
<th>Remarks</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOLEM</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A large number of non-disruptive discharges have been given as input to the neural network. For testing the ability of the ANN (i.e., how it triggers the disruption alarm), the probability of acquiring a disruption threshold needs to be found. The ANN model predicts the disruptions when it reaches the threshold value.

From the aforementioned point of view, the prediction of plasma discharge has been made successfully, and the result of the reduced network was compared with the original network. As a consequence, the probability of prediction accuracy of a TPR was three times greater than the false alarm rate, which indicates the minimization of false alarms.

Although there is a mathematical relationship between Murakami scaling and Hugill density limit [73], it has not been verified in this work.

### 3.2.5.2 Long Short Term Memory

Using one of the learning networks called LSTM [74], the disruptions have been predicted in ADITYA. On three grounds, this research study slightly varies from the others:

(i) The size of this tokamak is relatively very small, hence the disruptions have been predicted 12 ms in advance, unlike the big tokamaks like ASDEX or JET, where the disruptions have been predicted 100 ms in advance.

(ii) Different numbers of shots were used for training and testing in this study, whereas the existing model [75] used only one shot for training and three shots for testing.

(iii) The signal computation has been performed within a millisecond by using the proposed technique.

This network used a dataset with 119 disruptive shots, of which 83 were used for training and 36 for testing. By taking the few diagnostic signals as an input, the training and testing of the network have been conducted, and the prediction was done early in 12 ms. The best accuracy rate has been achieved for both testing and training is 94.58% and 95.39%, respectively. However, it was not enough to determine how many disruptions were correctly predicted.

The results are shown in Table 9 based on the time interval. Although it requires minimum time for computation, and the computational cost was very low, it requires large datasets for the network to learn different kinds of input, and also the number of premature alarms was very high.

<table>
<thead>
<tr>
<th>Types of Alarm</th>
<th>Prediction Occur based on Time Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>False Alarm</td>
<td>Greater than 40 ms</td>
</tr>
<tr>
<td>Premature Alarm</td>
<td>40-20 ms</td>
</tr>
<tr>
<td>Missed Alarm</td>
<td>After 8 ms</td>
</tr>
<tr>
<td>True Alarm</td>
<td>Rest of the alarm</td>
</tr>
</tbody>
</table>

### 3.2.5.3 Sequence Learning-based prediction using Long Short-Term Memory

To improve the prediction performance, the researchers proposed the same time-sequence LSTM technique [76] that was already used in the previous paper [74]. But the main difference was that, instead of taking 119 disruptive shots as an input, it takes a total of 125 shots with a combination of disruptive and non-disruptive. The following Table 10 represents the number of shots for training and testing purposes:
Table 10: Shots Allocation

<table>
<thead>
<tr>
<th>Number of Shots</th>
<th>Total Number of Shots</th>
</tr>
</thead>
<tbody>
<tr>
<td>125 shots</td>
<td></td>
</tr>
<tr>
<td>Training purpose</td>
<td></td>
</tr>
<tr>
<td>83 shots</td>
<td></td>
</tr>
<tr>
<td>Testing purpose</td>
<td></td>
</tr>
<tr>
<td>42 shots</td>
<td>(36 disruptive, 6 non-disruptive)</td>
</tr>
</tbody>
</table>

With the help of this dataset, the network model has been trained perfectly by taking a few diagnostic signals as input. The important note was that each of the signals must be pre-processed, which includes the following steps: (i) re-sampling and (ii) normalization, before being given any data as an input. It predicts the output value, whose range lies between 0 and 1. Suppose, if the output value was greater than the threshold, then it would indicate some major disruptions might occur. The prediction rate has been achieved in advance at 7-20 ms with an accuracy of 89%. But one major problem was that the non-disruptive shots were not used for training purposes due to their insufficient availability. For future work, many varieties of shots and diagnostic signals may be included to improve the performance of the model.

Table 11: Summary of data-driven approaches for ADITYA tokamak.

<table>
<thead>
<tr>
<th>Tokamak Reactor</th>
<th>Data-Driven approaches</th>
<th>Remarks</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADITYA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4 Discussions & Conclusion
This paper summarizes earlier work on plasma disruption forecasting using traditional and data-driven approaches. For different sorts of applications, each approach offers benefits and drawbacks. Due to the precision of their forecasting models' performance, data-driven approaches have attracted the most interest from researchers of all stripes. The creation of traditional techniques is straightforward but they only use the unprocessed raw data from a single signal to make predictions. So, it might not be able to make this forecast with precision. Therefore, data-driven procedures are thought to be the best approach and are able to offer higher predicting performance since they make use of many diverse diagnostic signals and normalize them before training. As a result, the output of the prediction model was far better than the classical model. In this study, we perform a survey encompassing various aspects of data-driven strategies on several experimental tokamak reactors. Based on the analysis, this review recommends numerous prospective future opportunities for research developments.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work has been financially supported by the Department of Science and Technology - Science and Engineering Research Board (DST-SERB) under Core Research Grant (CRG) Scheme (CRG/2021/003446).

References


51. BURRELL, HK. Overview of recent experimental results from the DIII-D advanced tokamak program. General Atomics, San Diego, CA (United States), (2002).


